

DEPARTMENT OF COMPUTER SCIENCE

MPhil Degree Oral Presentation

MPhil Candidate:	Ms Yichao ZHOU
Date	14 Aug 2018 (Tue)
Time:	2:30 pm - 4:30 pm (35 mins presentation and 15 mins Q & A)
Venue:	FSC703, Fong Shu Chuen Library, HSH Campus

“Lip Password-based Speaker Verification System with Unknown Language Alphabet”

Abstract

The traditional security systems that verify the identity of users based on password usually face the risk of leaking the password contents. To solve this problem, biometrics such as the face, iris, and fingerprint, begin to be widely used in verifying the identity of people. However, these biometric cannot be changed if the database is hacked. What's more, verification systems based on the traditional biometrics might be cheated by fake fingerprint or the photo.

Liu and Cheung (Liu and Cheung 2014) have recently initiated the concept of lip password, which is composed of a password embedded in the lip movement and the underlying characteristics of lip motion. Subsequently, a lip password-based system for visual speaker verification has been developed. Such a system is able to detect a target speaker saying the wrong password or an impostor who knows the correct password. That is, only a target user speaking correct password can be accepted by the system. Nevertheless, it recognizes the lip password based on a lip-reading algorithm, which needs to know the language alphabet of the password in advance, which may limit its applications.

To tackle this problem, in this thesis, we study the lip password-based visual speaker verification system with unknown language alphabet. First, we propose a method to verify the lip password based on the key frames of lip movement instead of recognizing the individual password elements, such that the lip password verification process can be made without knowing the password alphabet beforehand. To detect these key frames, we extract the lip contours and detect the interest intervals where the lip contours have significant variations. Moreover, in order to avoid accurate alignment of feature sequences or detection on mouth status which is computationally expensive, we design a novel overlapping subsequence matching approach to encode the information in lip passwords in the system. This technique works by sampling the feature sequences extracted from lip videos into overlapping subsequences and matching them individually. All the log-likelihood of each subsequence form the final feature of the sequence and are verified by the Euclidean distance to positive sample centers. We evaluate the proposed two methods on a database that contains totally 8 kinds of lip passwords including English digits and Chinese phrases. Experimental results show the superiority of the proposed methods for visual speaker verification.

Next, we propose a novel visual speaker verification approach based on diagonal-like pooling and pyramid structure of lips. We take advantage of the diagonal structure of sparse representation to preserve the temporal order of lip sequences by employ a diagonal-like mask in pooling stage and build a pyramid spatiotemporal features containing the structural characteristic under lip password. This approach eliminates the requirement of segmenting the lip-password into words or visemes. Consequently, the lip password with any language can be used for visual speaker verification. Experiments show the efficacy of the proposed approach compared with the state-of-the-art ones.

Additionally, to further evaluate the system, we also develop a prototype of the lip password-based visual speaker verification. The prototype has a Graphical User Interface (GUI) that make users easy to access.

***** ALL INTERESTED ARE WELCOME *****